On Firm Growth in Networks

Riccaboni, Massimo and Pammolli, Fabio
IMT Institute For Advanced Studies, Lucca

04. October 2001

On Firm Growth in Networks

Submitted to Research Policy, Special Issue on the Nelson and Winter Conference
September 15, 2001

Massimo Riccaboni* and Fabio Pammolli*

* University of Siena, Italy, mriccaboni@unisi.it

* University of Siena, Italy, pammolli@unisi.it

This research was supported by a grant from the Merck Foundation (EPRIS Project). We are grateful to audiences at Stanford University, University of Venice, and the Nelson and Winter Conference at DRUID for useful feedback. We thank Wesley Cohen, Steve Klepper, Lucia Cusmano, and Guido Caldarelli for helpful comments.

1. Introduction

Division of innovative labor and R&D collaborative contractual relationships are recognized as increasingly important economic phenomena (see Arrow, 1983; Arora, Fosfuri, and Gambardella, 2001).

In particular, networks of contractual relationships among firms specialized in research and exploration (*Originators*) and firms focused on development, production, and commercialization (*Developers*) are ever-widening organizational forms, especially in high-tech, knowledge-intensive fields (see Orsenigo, Pammolli, and Riccaboni, 2001).

In the last ten years, several studies have shown that network structure and positions in networks influence firm performance and growth (see Powell, Koput, and Smith-Doerr, 1996; Powell and Smith-Doerr, 1999) and, ultimately, market structure (Mc Lean and Padgett, 1996; Pammolli and Riccaboni, 2001). Moreover, most of the literature agrees that networks have to be analyzed as a distinct organizational solution for the access to outside knowledge sources, the coordination of heterogeneous learning processes by agents endowed by different skills, competencies, access to innovation, and assets (Pavitt, 2001).

In spite of growing consensus on networks as a distinct organization form and on their importance in processes of learning and evolution, economic models of division of (innovative) labor tend to focus on dyadic contractual relationships and on trade-offs defined at that level (Arrow, 1974; Williamson, 1991): particularly the tradeoff between economies from specialization exploited through task partitioning and the transaction costs involved in transferring knowledge and technological information through arm's length contracts (see Pisano, 1990; Teece, 1988; Teece, 1998; Arora, Fosfuri, Gambardella, 2001).

Against this background, the most relevant motivation for our analysis is an intellectual challenge to our understanding of the nature of '*networks of innovators*' (Freeman, 1991) and of processes of firm growth, which is revealed by a somehow unusual problem in the match between available theories and data. In fact, our empirical investigations on processes of firm growth in networks offer a neat picture, revealing the existence of scaling phenomena, with the firm connectivity distribution being well described by a power law of the form $N=kS^{-\alpha}$, where N is the number of firms with connections greater than S.

This result, which is stunningly equivalent to well known empirical regularities on processes of growth in several domains of both natural and social sciences (see Simon, 1955; Albert and Barabasi, 2001; Fujita, Krugman, Venables, 2000), is hard to reproduce in any of the available economic theoretical frameworks. It would be obviously possible to find parameter combinations in models that produce a good fit with real data. However, any explanation of such an apparently general phenomenon ought to be central in any modeling effort.

Along this way, we focus on the mechanisms behind the dynamic properties of growing networks and on firm growth in networks, unraveling striking analogies between processes of internal growth and processes of external growth through collaborative agreements.

We consider the links between *Originators* and *Developers* as instances of firm external growth. On the one side, *Originators* discover new technological opportunities and establish contractual relationships that generate income and give access to relevant assets. On the other side, *Developers* rely on collaborations with Originators to get access to outside knowledge sources and capture new technological opportunities.

We represent size and growth in terms of the number of connections of a given firm, seen as independent business opportunities of size unity arising over time (see Ijiri and Simon, 1977).

Moreover, we do show that the scale free behavior detected in networks can be accounted for by a very general and simple model, which is rooted in the 'old' stochastic approach to the analysis of firm growth.

We show that networks growth are shaped by entry of new firms and by proportional growth of the connectivity of individual firms, with remarkable departures from a regime of universal random growth.

In addition, different regimes of growth are found to be at work for Originators vs. Developers, reflecting differences in the processes of generation and absorption/development of technological opportunities.

In particular, the population of Originators is characterized by a regime of proportional growth which corresponds to a 'popularity is attractive' mechanism (see also Zucker, Darby, Brewer, 1997), while for Developers this mechanism is attenuated by a random component.

While this result cannot be fully explained given the present status of our knowledge, it is coherent with an interpretation of firm growth and networking activities which is rooted in a competence-based view of organizational growth and division of labor (Penrose, 1995; Richardson, 1972; Nelson and Winter, 1982; Dosi, 2000).

In particular, the empirical findings presented in this paper, as well in Orsenigo, Pammolli, Riccaboni, 2001 and in Pammolli and Riccaboni, 2001) show that processes of network growth are sustained by the existence of dynamic complementarities between patterns of specialization in knowledge production (Originators) and processes of diversification of in-house capabilities by large multi product, multi technological companies (Developers) (see also Granstrand, Patel, Pavitt, 1997; Pavitt, 2001).

We do think that our analysis points to some basic principles behind the growth of firms in technological networks, providing a simple benchmark to be used in the context of future investigations.

In addition, one important feature of our work is related to the fact that, since we are dealing with the dynamics of a set of links, we can exploit the duality of the overall system, extracting topological information which can be used to uncover the underlying causal data generating mechanisms. That is to say, one important contribution of the analysis of firm growth in systems of division of labor can be the possibility to produce plausible restrictions on the acceptable classes of conditional predictive distributions and on the dynamics of the stochastic processes which generated them, so contributing to a better understanding of firm growth in general.

2. Firm Growth and Connectivity in Networks

It is our claim that firm growth in networks can be fruitfully framed in terms of the 'old' stochastic tradition in the analysis of processes of firm internal growth, with particular reference to the seminal contributions of Simon and colleagues (see Simon, 1955; Ijiri and Simon, 1977).

Both business size distributions and nodes degree distributions of many real-world networks exhibit heavy tails and power-law scaling of the form $P(k) \sim k^{-\gamma}$ (see Sutton, 1997; Brock, 1999; Albert and Barabasi, 2001). The connectivity distributions of networks with complex topologies such as the world wide web, the internet, phone call and power networks, the movie actor collaboration network, the science collaboration graph, the web of human sexual contacts, the citation network of scientists, follow scale-free power laws, reflecting some major departures from a regime of 'universal' random growth (Barabási and Albert, 1999).

To make a long story short (see Riccaboni, 2000), the origin of scale-free behaviors in networks can been accounted for by a simple model for scaling in growth processes that was proposed by Herbert Simon (1955), in order to give an interpretation of distributions such as word frequencies in texts or population figures of cities.

Simon models the dynamics of a system of elements with associated counters (business opportunities of size unity) where the dynamics of the system is based on constant growth via the addition of new elements (new business opportunities) as well as incrementing the counters at a rate proportional to their current values. First, networks can grow by the addition of new nodes that become linked to existing ones. Second, networks growth can be driven by a popularity mechanism (preferential attachment).

Interestingly enough, these two mechanisms can be considered as particular instances of the model which was solved by Simon in his 1955 paper, in which the Pareto distribution is derived from “simple and economically plausible assumptions”, namely size independence of percentage growth rate (the Gibrat’s law of proportionate effect), and constancy of the entry rate. In particular, the original Simon model accounts for a robust empirical regularity that has been detected in many networks across different fields, irrespectively of their nature and components: that is, the probability distribution of the number k of links that point to a particular node (i.e. web page, scientist), P(k), decays following a power law $P(k) \sim k^{-\gamma}$, with the scaling exponent γ being very close to 2, both for the distribution of in coming and out coming links (for a review, see Albert and Barabási, 2001).

Given the pervasiveness of scale free distributions across different empirical domains, we retain here the Pareto curve in our analysis of firm growth in networks. .

In particular, Ijiri and Simon have shown, in the case of business firm size, the existence of systematic departures from Pareto. Equivalently, most empirical connectivity distributions in networks depict similar flattened upper tails, suggesting equivalent departures from the Pareto law, with nodes with a low connectivity following a different distributional model (possibly Poisson, or a combination of Poisson and power law).

In the next session, we will show that the mechanisms identified by Simon and colleagues to explain the observed departures from the Pareto size distribution - namely, M&As and growth autocorrelation- hold also in the case of firm growth in networks: (a) The probability of a firm disappearing is not independent of its connectivity: poorly connected organizations evidence a higher probability to be acquired by core players than vice versa; (b) The growth of firms within the network is characterized by autocorrelation, which tends to vanish over time: recent links have only a short-run effect upon firm’s future probability of relinking (that is, the probability of establishing a new link conditional upon having already established a link); the effect of a given opportunity (deal established) on future collaborations decays as time goes by.

Despite the general validity of the framework that we have recalled, we detect the existence of systematic differences between the curvatures of the connectivity distributions of Originators vs. Developers.

In the next session we will show, by means a simple simulative model, that these differences can be considered the result of inherently different mechanisms of growth at work for the two types of firms.

3. Firm Growth in Networks: Empirical and Simulative Results

In this section we refer to a domain, biopharmaceuticals, in which the R&D network among firms has grown substantially in the last 30 years.

New bodies of knowledge have generated a plethora of scientific and technological opportunities, nurturing a continuous flow of entry of new firms, as well as an extensive division of innovative labor between *Originators* and *Developers* of R&D projects (Orsenigo, Pammolli, and Riccaboni, 2001).

The expansion of the network has been driven by the entry of new firms and by the addition of new collaborations. In particular, during the Nineties, the number of research alliances has grown fourfold, while the number of firms has almost doubled. At the same time, the number of M&A events has been steadily high, culminating with a few mega-mergers in the last few years[1] (see Pammolli and Riccaboni, 2001).

Data used for this study are drawn from the Pharmaceutical Industry Database (PHID) at the University of Siena.
An important feature of PHID is that it provides information on typology, technological content, and date of signing for 5353 collaborative agreements and 989 mergers and acquisitions (M&As), involving 1583 firms worldwide.

---

[1] 1996: Ciba-Geigy – Sandoz (*Novartis*); 1997: Roche – Boehringer Mannheim; 1998: Hoechst Marion Roussel – Rhône-Poulenc Rorer (*Aventis*); Sanofi – Syntélabo; Astra – Zeneca (*AstraZeneca*); 1999: Pharmacia & Upjohn – Monsanto (Pharmacia Corp.); 2000: Glaxo Wellcome – SmithKline Beecham (*Glaxo SmithKline*); Warner Lambert – Pfizer.

Here, given our focus on division of innovative labor– we have selected 3807 R&D collaborative transactions subscribed by 349 pharmaceutical companies, and 1100 Dedicated Biotechnology Firms (DBFs). For each firm we have collected additional information on location, size, main areas of activity, age, and type.

For each contract, we have recorded the following *transaction-specific attributes*:

Date of signing;

Stage of project development at subscription (i.e. discovery, preclinical, clinical);

Technological content (i.e. gene therapy, genomics, molecular diversity…);

Targeted diseases (i.e. AIDS, Alzheimer, Cancer…);

Typology (viz. license, joint venture, co-development…).

For 3171 R&D collaborative agreements we are able to distinguish an *Originator* (Licensor) from one or more *Developers* (Licensees).

Based on our data set, we begin our investigation by looking at the connectivity distributions for Originators and Developers. In Figure 1 the integrated connectivity density distribution for both Originators and Developers is plotted on a double log scale. The upper tails of both distributions are well fitted by a power law with exponent –1, correspondent to $\gamma=2\pm0.1$. Figure 1 reveals, however, the existence of a remarkable concavity of the actual connectivity distributions, which substantiates a rather significant departure from the theoretical power law distribution.

[FIGURE 1 ABOUT HERE]

In order to investigate the mechanisms behind the observed departures, we sort firms in decreasing order of connectivity and plot the relationships between rank and connectivity on a log-log scale. The results are shown in Figure 2 and 3.

As noticed above, a first economic mechanism that sways the degree distribution form Pareto is the process of consolidation. Since in the pharmaceutical industry merged and acquired firms usually remain separated *de facto* for long, we are allowed to follow their growth processes even after tie-up events and to evaluate the degree distribution of single freestanding *divisions* as if they never collapse into they relative *holdings*.

The comparison of distributions in Figure 2 reveals that a significant fraction of the departure from the Pareto connectivity-rank distribution (straight line on a double-log scale) can be ascribed to M&As. As it is evident, the slope of the post-merger distribution (*holdings*) is steeper than the slope of the "mergerless" distribution (see also the OLS estimates in Table 1 below). Moreover, the concavity of the distribution is influenced by M&As. The deviation of the post-merger from the Pareto distributions is wider than the correspondent departure from the "mergerless" distribution. Not surprisingly, the probability of a firm disappearing is not independent of connectivity, since less connected organizations have a higher probability to be acquired by core players than vice versa.

[FIGURE 2 ABOUT HERE]

So far, we have discussed the effects of M&As on the connectivity distribution. However, the consolidation process accounts for only a fraction of the observed departure form the Pareto distribution. Autocorrelation of growth opportunities is a second possible cause leading to the concavity of curve. In order to test if this second mechanism holds in the case of firm growth in networks, we now focus on the deal-making activity between standing-alone firms.

As it is evident from Figure 3, both mergerless distributions show remarkable departures from Pareto. Figure 4 shows that the probability of capturing a new opportunity decays in time. As noticed by Ijiri and Simon, 1977, this mechanism substantiates can account for the observed deviation from Pareto.

The above results are confirmed in Table 1, in which we perform an OLS estimate of the theoretical Pareto distribution: $\log k = \log A + b \log r + e$. The connectivity of each firm (k) and its rank (r) are shown to be linearly related on a double-log scale. Table 1 reports also the values of the intercepts (log A, i.e. the logarithm of the largest firm' connectivity) and the estimated slope coefficients (*b*). In order to measure the extent of the concavity we have added a cubic term and compare the estimated coefficients (*c*).

However, Figure 3 shows also systematic differences between the curvatures of the connectivity distributions of Originators vs. Developers. These findings are suggestive of different relational behaviors for the two types of firms, which are not explained by differences in intensity and decay of growth autocorrelation, which appear to be very similar by looking at Figure 4.

[FIGURE 3 ABOUT HERE]

[TABLE 1 ABOUT HERE]

[FIGURE 4 ABOUT HERE]

In order to improve our understanding of these different patterns of firm growth in networks, we introduce a simulative model, which is based on two simple parameters and make possible a better characterization of the processes of growth for Originators and Developers.

The model is based on the simple assumptions of entry and proportional growth. A parameter (p) sets the interdependence between the growth of the network and the flows of firm entry. A parameter (q) gives the probability of having a cumulative relational regime, reflecting the relative growth of number of nodes versus number of links.

Each simulation starts with N isolated nodes (firms).

At each time step, with probability p, a new Originator enters the network, whilst with probability (1-p) a link is originated by an already active firm.

With probability q, an Originator links to a Developer chosen preferentially, in proportion to its connectivity. In this case a new link is drawn from an Originator to a

Developer, which is selected with probability Π(d) proportional to its degree k(d). Based on the evidence discussed in Powell, Koput, and Smith-Doerr, 1996, as well as in Orsenigo, Pammolli, and Riccaboni, 2001, we model Π(d) as a linear function of k(d). With probability (1-q), an Originator establishes a new link with a preexisting Developer chosen at random.

We tested different versions of the simulative model for different combinations of p and q. In a nutshell, two different generative processes turn out to be in place for Originators and Developers. In the case of Originators, the actual connectivity distribution is accounted for by a regime of preferential attachment and sustained entry (p=.5, q=1).

On the contrary, in the case of Developers the simulative model that better approximates the real-world distribution is a mixture of the random and the cumulative generative processes, with sustained entry (p=.5, q=.5)..

4. Concluding Discussion

In this paper we have shown that a simple generalization of stochastic explanations of internal firm size and growth fit a whole range of empirical findings.

Along this line, we have introduced a model, which can be used as a benchmark in the analysis of firm external growth in networks of collaborative agreements.

In the context of a specific industry, we have shown that the growth of the overall network of R&D collaborative agreements is shaped by the interplay among a differentiated set of regimes of firm growth, with a crucial role being played by entry and by a mechanism of proportional growth.

The scale-free structures that we have found to be in place in the network of collaborative agreements in pharmaceutical R&D can be considered as one striking outcome of a fairly general 'popularity is attractive' principle, which seems to sustain also the growth of systems of division of labor and of firms acting within them.

Being very general, mechanisms behind external growth in networks do not seem to differ from the ones that sustain firm internal growth.

We do think that this result is deeply suggestive of the existence of organizational principles that are general in nature, and map on both the internal structure of firms and the structure of markets and networks.

Moreover, we have shown that the dual nature of networks can convey information on topological properties of industries and roles/positions of firms within them (to begin with, the distinction between Originators and Developers), which can be used to understand some fundamental structures, mechanisms, and generative processes behind the growth of firms and industries, in the direction of building parsimonious and, at the same time, realistic, representations.

At present, our analysis has some obvious limitations. First, apart from information on firms' age and on the distinction between Originators and Developers, we did not take into account any node-specific attribute. Second, we have considered links of size unity, without addressing the properties of weighted networks and interactions strength. Third, the relational propensities of different nodes stay unchanged in our model. Finally, the exit of nodes from the network was not accounted for.

These shortcomings notwithstanding, this paper should be considered as a first step towards the understanding of some general determinants of firm growth in networks. Despite its limitations, it provides a parsimonious and general framework to 'reverse engineering' the growth of networks in different industries, as we attempt to make our models more realistic.

Some of the current limitations of our analysis could be overcome, in the future, based on a higher availability of data on real systems and, in particular, of detailed topological and economic information on real-world networks.

While at present such data are relatively rare, the increasing interest in industrial networks is leading to the development of suitable data sets, offering further guidance for modeling and interpreting the growth of these complex and important economic systems.

References

Albert, R., Barabási, A. L., 2001, *Statistical mechanics of complex networks*, Department of Physics, University of Notre-Dame, Notre-Dame, In, cond-mat/010696v1.

Arora, A., Fosfuri, A., Gambardella, A., 2001, *Markets for technology: The economics of innovation and corporate strategy*, MIT Press, forthcoming.

Arora A., Gambardella, A., 1994, "The changing technology of technological change: General and abstract knowledge and the division of innovative labor", *Research Policy*, vol. 23, pp. 523-532.

Arrow, K. J., 1974, *The limits of organization*, Norton, New York.

Arrow, K. J., 1983, "Innovation in large and small firms", in Ronen, J., ed., *Entrepreneurship*, Lexington Books, Lexington, Ma.

Barabási, A.L., Albert R., 1999, "The emergence of scaling in random networks", *Science*, vol. 286, pp. 509-512.

Brock, W. A., 1999, "Scaling in economics: A reader's guide", *Industrial and Corporate Change*, vol. 8, n. 3, pp. 409-446.

Dosi, G., 2000, *Innovation, organization, and economic dynamics*, Edward Elgar, Cheltenham.

Fujita, M., Krugman, P., Venables, A. J., 1999, *The spatial economy. Cities, regions, and international trade*, MIT Press, Cambridge, Ma.

Granstrand, O., Patel, P., Pavitt, K., 1997, "Multi-technology corporations: Why they have 'distributed' rather than 'distinctive core' competencies", *California Management Review*, vol. 39, pp. 8-25.

Ijiri, Y., Simon, H.A., 1977, *Skew distributions and the sizes of business firms*, North Holland, Amsterdam.

Malerba, F., Orsenigo, L., 2000, "Knowledge, innovative activities, and industrial evolution", *Industrial and Corporate Change*, n.2, pp. 289-313.

Mc Lean, P., Padgett, 1996, "Was Florence a perfectly competitive market? Transactional evidence from the Renaissance", *mimeo*.

Nelson, R., Winter, S., 1982, *An evolutionary theory of economic change*, Harvard University Press, Cambridge, Ma.

Orsenigo, L., Pammolli, F., Riccaboni, M., 2001, "Technological change and network dynamics. Lessons from the pharmaceutical industry", *Research Policy*, vol. 35, n. 3.

Pammolli, F., Riccaboni, M., 2001, "Technological regimes and the growth of networks. An empirical analysis", *Small Business Economics*, forthcoming.

Pavitt, K., 2000, "Can large Penrosian firms cope with the dynamics of technology?", *SPRU, Working Paper n. 68.*

Penrose, E., 1995, *The theory of the growth of the firm*, Oxford University Press, Oxford.

Pisano, G., 1990, "The R&D boundaries of the firm: An empirical analysis", *Administrative Science Quarterly*, vol. 35, pp. 153-176.

Powell, W.W., Koput, K., Smith-Doerr, L., 1996, "Technological change and the locus of innovation: networks of learning in biotechnology", *Administrative Science Quarterly*, vol. 44, n. 1, pp. 116-145.

Riccaboni, M., 2000, "Network evolution in drug discovery", *Ph.D dissertation*, Sant'Anna School of Advanced Studies, Pisa.

Richardson, G., 1972, "The organization of industry", *Economic Journal*, vol. 82, pp. 883-896.

Simon, H.A., 1955, "On a class of skew distribution functions", *Biometrika*, vol. 42, pp. 425-440.

Sutton, J., 1997, "Gibrat's legacy", *Journal of Economic Literature*, Vol. XXXV, March, pp. 40-59.

Teece, D.J., 1988, "Technological change and the nature of the firm", in Dosi, G., et al., eds., *Technical change and economic theory*, Pinter, London.

Teece, D.J., 1998, "Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets", *California Management Review*, vol. 40, n. 3, pp. 55-79.

Williamson, O.E., 1991, “ Comparative economic organization: The analysis of discrete structural alternatives”, *Administrative Science Quarterly*, vol. 36, pp. 269-296.

Winter, S., 1984, “Schumpeterian competition in alternative technological regimes”, *Journal of Economic Behavior and Organization*, vol. 5, pp. 287-320.

Zucker, L., Darby, M., Brewer, M., 1997, “Intellectual human capital and the birth of US biotechnology enterprises”, *American Economic Review*, vol. 87, n. 3.

Figure 1- Integrated Connectivity Density Distribution: Originators and Developers

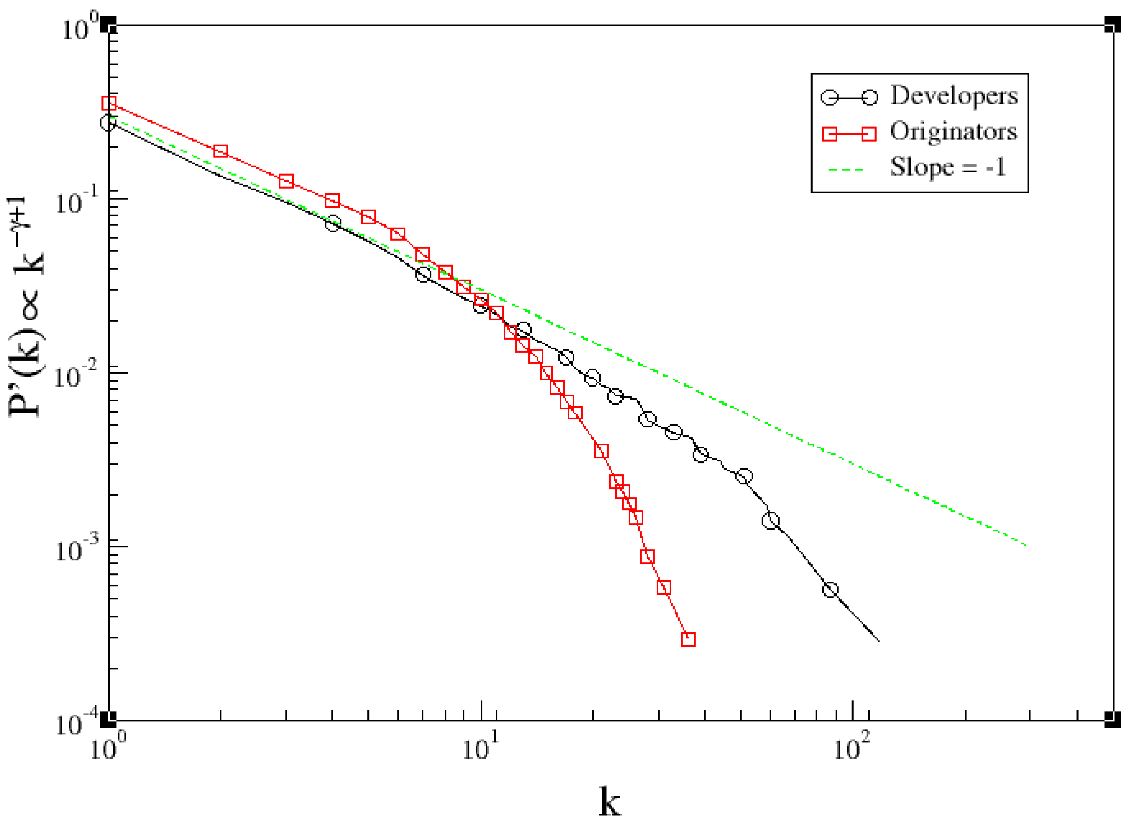


Figure 2- Rank-Size Distribution of Firms in the Network of Collaborative Agreements: The Effect of M&As

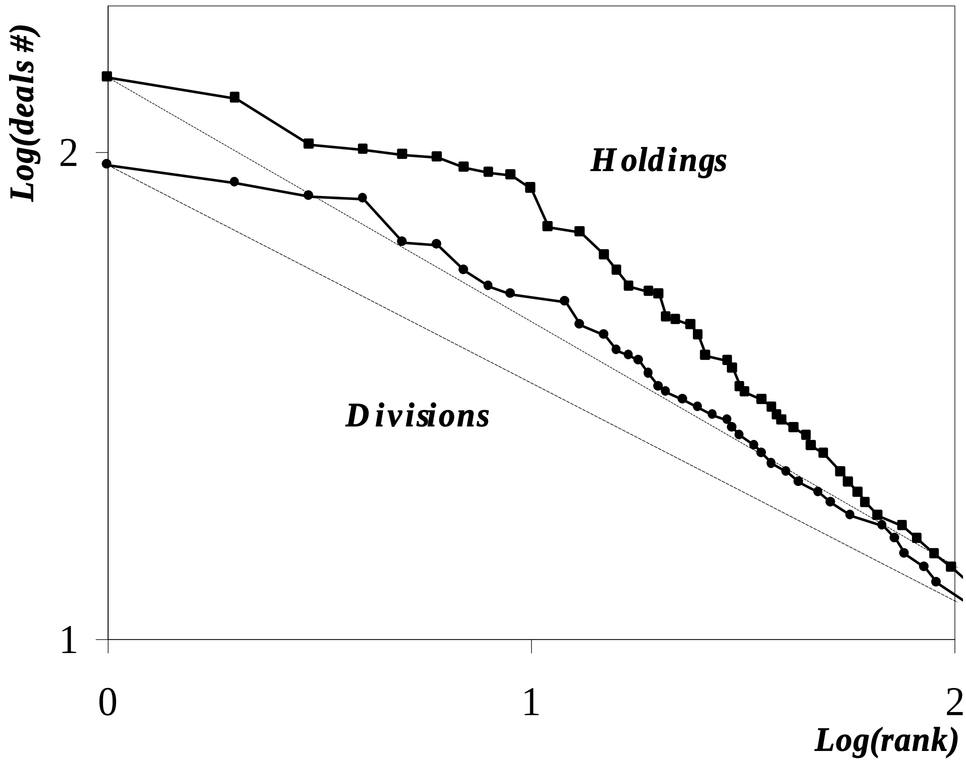


Figure 3- Rank-Size Distribution of Firms in the Network of Collaborative Agreements: Originators and Developers

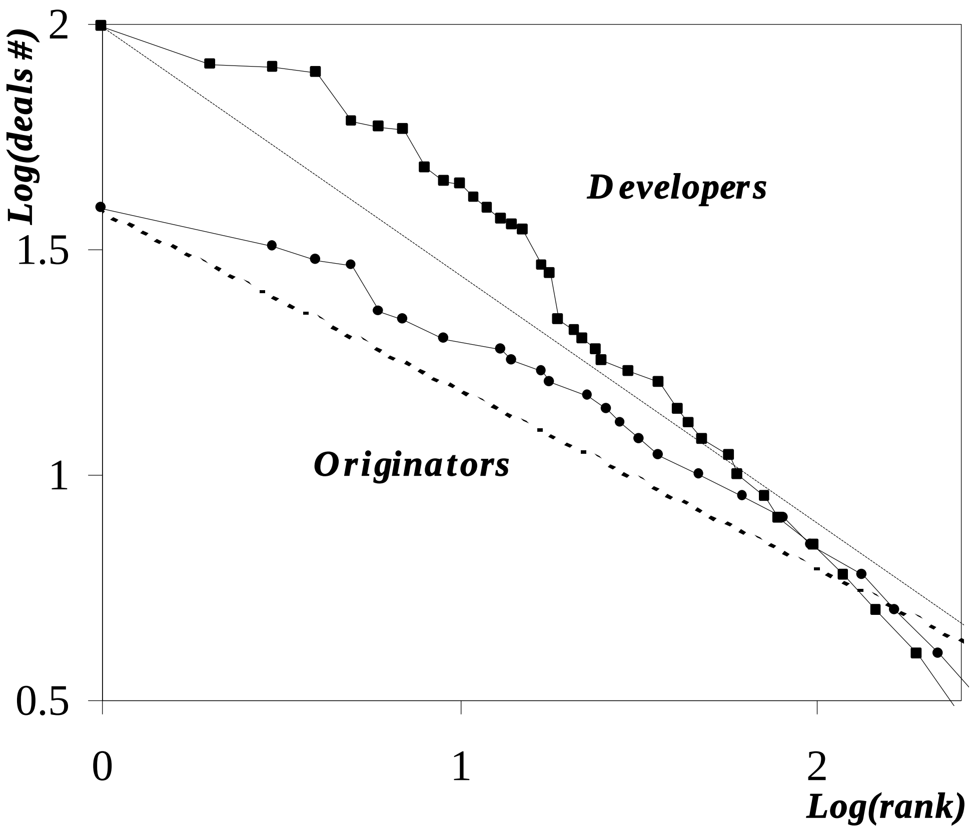
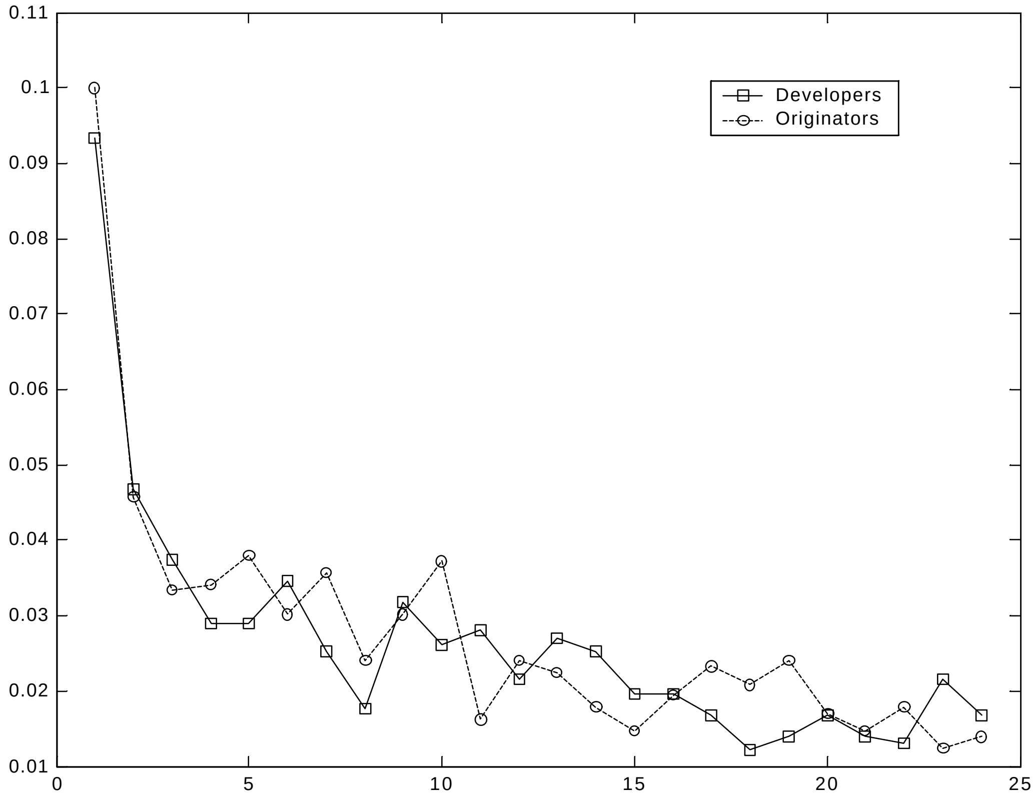


Table 1. Pareto Regressions

	log A	*b*	*c*	R^2
Holdings, linear	2.169	-0.500	-	0.84
Holdings, cubic	2.169	-0.117	*-0.197*	0.99
Divisions, linear	1.973	-0.446	-	0.85
Divisions, cubic	1.973	-0.126	*-0.158*	0.99
Developers, linear	1.991	-0.551	-	0.88
Developers, cubic	1.991	-0.206	*-0.186*	0.99
Originators, linear	1.591	-0.400	-	0.87
Originators, cubic	1.591	-0.115	*-0.141*	0.99

Figure 4- Probability of Relinking at Different Time Steps: Originators and Developers

Probability of Relinking



Time (months)